10. Evaluate the Result

This chapter discusses the Partnership approach to developmental and operational test and evaluation.

In particular, this chapter covers the following topics:

- Understanding the military worth assessment process
- Understanding the key insights and redesign ideas
- The step-by-step process

Much of the process described in this chapter parallels the EW test process. For greater detail about the EW test process, please refer to Air Force Manual 99-112, Electronic Warfare Test and Evaluation Process—Direction and Methodology for EW Testing.

10.1 Understanding the Military Worth Assessment Process

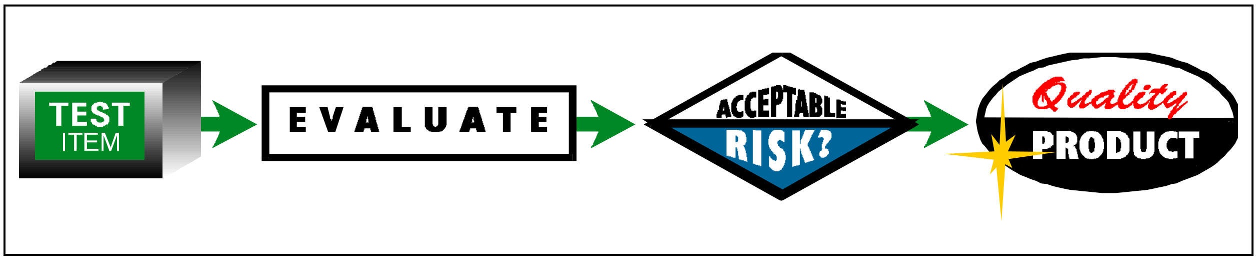

Figure 10-1. Evaluate the Result Process Flow. *During the evaluate the result activity, we gather information that helps us decide how the acquisition program ought to proceed.*

Figure 10-1 shows the EW test process. During this activity, we gather information that lets us determine how the acquisition program ought to proceed and identify system characteristics that provide valuable information to the warfighter.

10.1.1 Evaluating to Inform

We evaluate a developing solution throughout the acquisition process. All test and evaluation (T&E) activities aim to gather realistic and objective data, and perform impartial analysis. The operational test community has the additional responsibility of assessing the operational effectiveness and suitability of new systems for decision makers.

In this chapter, we define test in the broadest sense, to include all activities that gather data about a test item and that support decision making.

In particular, the purpose of test and evaluation is to:

- Reduce risk early in the acquisition process by discovering flaws in the product that would be very costly to fix later.
- Demonstrate system effectiveness and suitability that proves new and modified systems are being properly developed and improved, and will meet the needs of the user.
- Contribute timely, accurate, and affordable information to support life cycle acquisition and support choices of decision makers.

In addition to ensuring that all test and evaluation activities provide insight into the function of a solution, the Partnership's Military Worth Method entails some change in the way the test and evaluation community performs the evaluation.

To characterize the contribution of a system to mission objectives, tests and evaluations need to show how a system impacts the warfighter's mission plans—that is, tests and evaluations need to gather information about whether and how a system buys back airspace so that missions can be completed successfully.

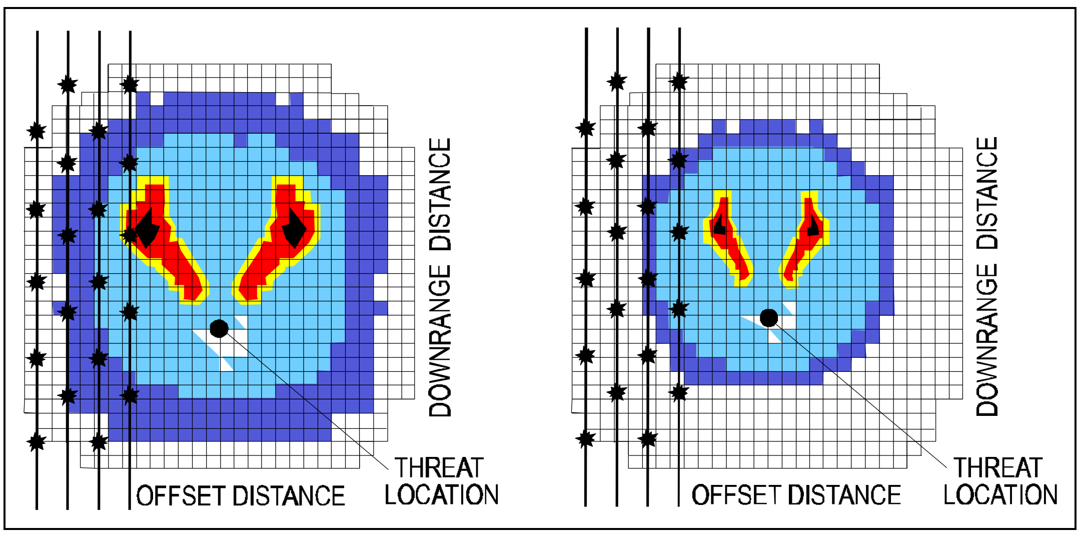

Figure 10-2. EW and Airspace Bought Back. *Test and evaluation needs to provide information about whether and how a system buys back airspace. Data about the location of each hit—the explosions on each flight path—is necessary for a military worth evaluation.*

10.1.2 Evaluating Results and the DoD 5000 Phases

Test and evaluation happens primarily during Phases I, II, and III of an acquisition. During Phase 0, Concept Exploration, test and evaluation activities are more commonly known as technology or proof of principle demonstrations.

In Phase I, contractors and the government typically test and evaluate prototypes of components, subsystems, or entire systems. In Phase II, the government often evaluates test items that are representative of production articles. In Phase III, the government conducts tests on production models.

The following table indicates the tasks that are specific to each phase. The principles we discuss in this chapter apply in every phase beyond Phase 0.

DoD 5000 Phase	Distinguishing Features of Each Phase
Phase 0: Concept Exploration	• Typically models of preferred concepts and critical technologies • Technology demonstrations • Data validates and refines modeling and simulation set • Supports narrowing the range of preferred concepts
Phase I: Program Definition and Risk Reduction	• Typically breadboards, brassboards, and/or prototypes • Much system integration lab (SIL), hardware in the loop (HITL), and some open-air range (OAR) • Data validates and refines modeling and simulation set • Heavy developmental test (DT), early operational test (OT) • Supports narrowing range of solutions and design trades

DoD 5000 Phase	Distinguishing Features of Each Phase
Phase II: Engineering and Manufacturing Development	• Typically pre-production models • Modeling and simulation bounds needed test data; data validates, refines modeling and simulation • All test environments • Heavy DT and OT • Supports production decision
Phase III: Production, Fielding/Deployment, and Operational Support	• Typically production models • Modeling and simulation bounds needed test data; high model maturity • DT for modifications and upgrades; continued operational T&E • Supports modification/upgrade decisions and tactics development

Figure 10-3. DoD 5000 Phases. *The specific activities entailed by test and evaluation vary depending on the phase of the acquisition, but the purpose of, and approach toward, development should be consistent.*

10.2 Understanding the Key Insights and Redesign Ideas

Key Insights and Redesign Ideas	• Military worth quantified and communicated • Development of greater modeling and simulation capabilities

10.2.1 Military Worth Quantified and Communicated

To gain the benefits promised by the Military Worth Method, we must ensure that the following principles apply throughout the EW test and evaluation process:

- We need to generate a detailed test plan that combines 1-v-1 effectiveness evaluation through modeling, simulation, and testing with the geometry-based campaign metrics.
- The test approach should employ geometry-based engagement-level performance metrics compatible with how the Military

Worth Method analyzes the impact of reduced threat lethality envelopes.

- Linkages should be established between the test environment and the operational scenario that was used to establish the requirements trade space.
- Data evaluation needs to consider hit offset and aspect angle information to support geometrically based analysis.

Making decisions based on military worth demands evaluations that quantify the value of the solution and relate it back to the warfighter's needs. Arriving at this quantified value and implementing it as the standard for EW test and evaluation may create some difficulties. However, a consistent and verifiable assessment of military worth, used throughout test and evaluation activities, will allow us to perform cost as an independent variable (CAIV) trade and get the best value for the warfighter.

A test and evaluation process that helps us assess the military worth of a solution can allow the warfighter to improve combat plans, and permits decision makers to rely on the analysis made available through modeling and simulation. Specific benefits of the new test and evaluation methodology include:

- A clear match between the requirements from the Operational Requirements Document (ORD) and the test and evaluation data from the test community.
- An ability to demonstrate key system and component performance and a means of linking performance to military worth.
- Operational test reports that help the warfighter perform tactical planning.

One of the benefits of the Military Worth Method is that it allows us to no longer test to reduction in lethality (RiL). According to Marion Williams, technical director for the Air Force Operational Test and Evaluation Center (AFOTEC), quoted in a recent issue of *Aviation Week and Space Technology*:

> *Current evaluation techniques simply do not measure the military worth of EW systems.*

The Military Worth Method, which uses reduction in low-kill offset as its test measure, will allow us to assess the military worth of EW systems.

EW Test and Evaluation and the Military Worth Framework

In addition to collecting data that allows us to characterize the geometric benefits of the solution, we need to analyze that data according to the framework used in the Partnership's Military Worth Method. As a result, the test and evaluation community must aggregate test data and show how it relates to the scenario we used to establish requirements. This relation allows decision makers to see how changes in the lower-level attributes of a system affect higher-level indicators of a system's performance and military worth.

Test and evaluation results provide us with insight into the first three levels of the military worth framework: technical attributes, operational functions, and operational capabilities of a system.

Using modeling and simulation tools and rigorous analysis, we can aggregate from these results to the higher levels of operational tasks, operational objectives, and campaign objectives. The primary measure of an EW system's military worth—how many additional Air Tasking Order (ATO) tasks it allows us to achieve—is the basis for decisions we make within the requirements trade space.

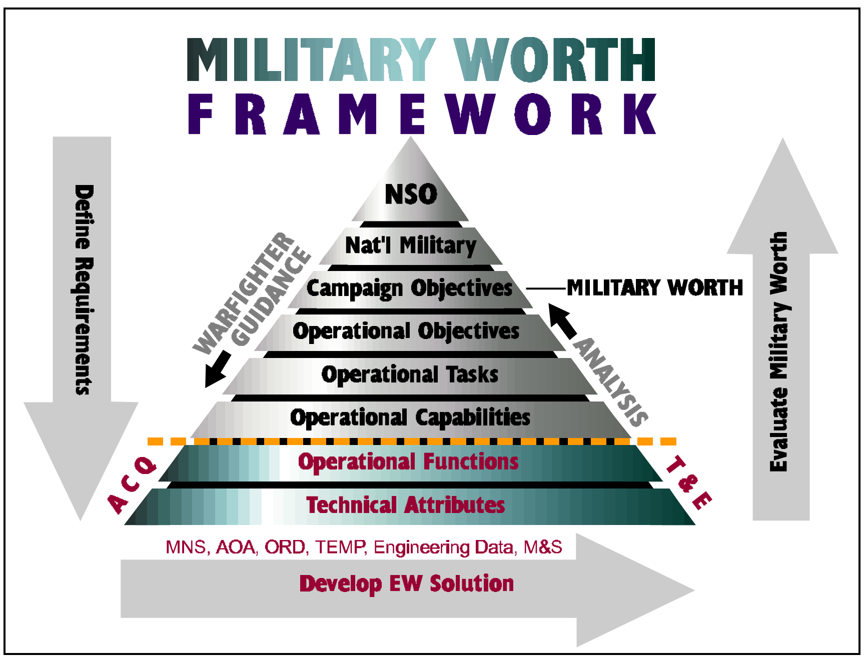

Figure 10-4. Using Test Results Within the Military Worth Framework. *The levels of technical attributes, operational functions, and operational capabilities are the focus of most test and evaluation efforts.*

Using of Modeling Tools to Determine Military Worth

Military worth cannot be measured directly. To quantify military worth and determine the effects of an EW system on campaign and mission objectives, we need to use modeling and simulation tools of the following levels of complexity:

- Campaign level simulations
- Mission level simulations
- 1-v-1 engagement simulations

Each of these simulations provides valuable information about the relationships between the threat, the planned mission, and the expected contribution of an EW system. The high-level analysis performed by campaign level simulations is fed into the mission level simulation, along with data about 1-v-1 engagements, and the mission level simulation calculates the effectiveness of the system in terms of ATO objectives achieved.

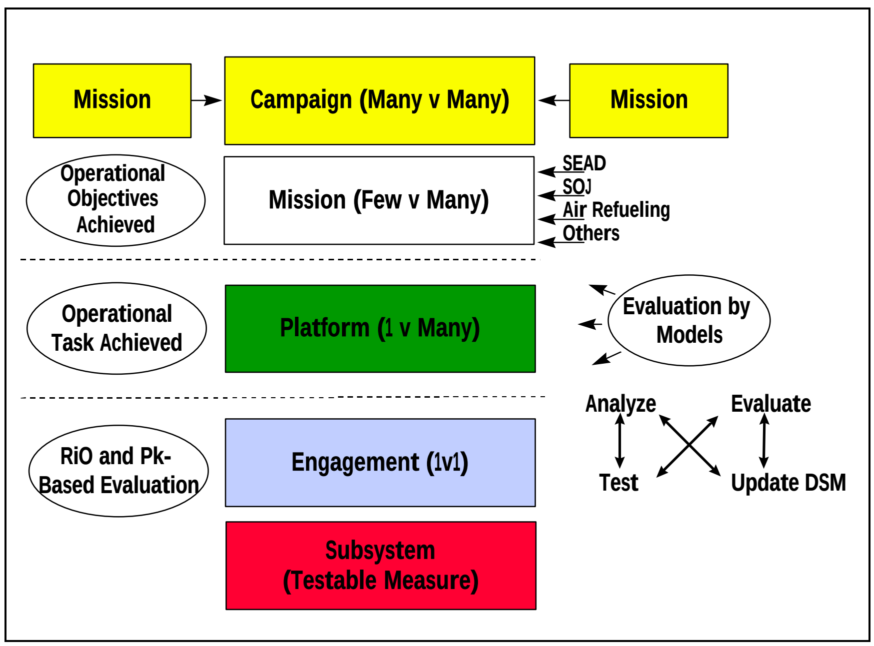

Figure 10-5. Relationship Between Modeling and Simulation Tools and the Test and Evaluation Process. *Simulations perform different levels of analysis to produce a calculation of targets at risk. In general, operational test and evaluation stops at the level of operational tasks achieved, and developmental test and evaluation stops at the level of RiO and* P_k*-based evaluation.*

Using the RiO Measure to Report Test Results at the Mission Level

Whatever type of test we perform—from early lab tests on system components or open-air range (OAR) tests—we need to ensure we continue to collect all the data we need to analyze the results at the level of mission effectiveness.

For a full discussion of RiO, see Section 4.3.3, The Role of Low-Kill Offsets and P_k Grids.

Data we collect through test and evaluation activities must help us understand not only whether a hit or miss occurred with the solution enabled but also where a hit occurred in relation to the threat. While this kind of data collection is clearly required during OAR tests, we also need to employ modeling tools that help us make these determinations throughout system development. The results of this analysis will allow the warfighter to plan strategies based on an accurate picture of the reduction in low-kill offset (RiO) provided by the solution.

One reason RiO was chosen as a measure is because the test community has the resources and abilities to implement this measurement immediately. While there may potentially be more accurate or sophisticated measures, the warfighters' needs are best served by an effective system that can be implemented now.

Conceptually, testing for RiO does not involve gathering data in a different manner nor does it involve gathering new data. The difference is that RiO, unlike RiL, uses all the data recorded, namely the spatial correlation information, which was not used in the evaluation.

Measuring RiO will require careful attention to detail so that digital system models (DSMs) can be updated accurately using the offset and aspect angle values we are computing from the test data. This information can then be aggregated through the modeling and simulation tools, which can calculate targets at risk for the measured RiO.

10.2.2 Develop Greater Modeling and Simulation Capabilities

Military worth is not testable by itself. The test and evaluation community must use modeling and simulation tools to apply results of testing to the military worth framework. Although models have always been a part of electronic warfare test and evaluation, they must play a greater role in the evaluation of test results. This means extending test results to the mission level.

The DSM should be developed to help in system design, perform pre-test analysis, and evaluate the results from testing.

The test and evaluation community has traditionally performed the following functions:

- Planned a test program based on ORD requirements
- Ensured the availability of the necessary supporting infrastructure
- Provided, maintained, and upgraded test environments
- Measured the results of 1-v-1 engagements

Once these tasks were complete, test results were compared to ORD values and a pass/fail evaluation was performed.

Using the Partnership Process, testers also conduct engagement level measurements. However, the data gathered from these tests is not directly compared to ORD values, but is used instead to update the solution's DSM. The DSM is currently used in ESAMS to build a new P_k grid, which represents the solution's capability in each engagement scenario.

These P_k grids are then incorporated by mission level and campaign level simulations to determine the overall operational performance achieved by the system. This value is passed on to the decision makers (along with cost and schedule information), who determine whether the current performance is worth the investment, and what adjustments, if any, should be made to the acquisition program.

Relating Test Results to Modeling and Simulation

We gather test data to learn whether our current P_k grid prediction for a system's performance against each threat is correct. What we test, at the component or overall system level, is whether selected points match our P_k grid prediction.

In the later stages of test and evaluation, it would be too costly and time consuming to test every point on a P_k grid, so these efforts focus on gathering data for those points where we have lower confidence in our predictions and at other selected points to verify that regions of high confidence are indeed accurate.

Analysis of Pk grids using test results should proceed according to the following method:

1. Update DSM for the system under test and use these in a 1-v-1 engagement model (such as ESAMS).
2. Evaluate the prediction to determine test points. To do this properly, we need to understand the areas of our prediction that

present high and low confidence. The determination of these points is normally part of the detailed technical analysis performed by both government and contractor personnel.

3. Obtain test data correlating to certain points in the predicted P_k grid:

   — If the values are different in a certain area, then we need to understand why the prediction is different from what we have measured.

   — Based on that understanding, change the DSM if the difference was because the DSM incorrectly represented the performance of the system

4. Use updated DSM to generate new P_k grid.

5. Use the new predictions in SUPRESSOR or equivalent mission level model to generate new assessments of the military worth of the system in terms of ATO tasks accomplished (for example, targets at risk).

Using the targets at risk (TAR) measurement, decision makers (the warfighter or program manager) determine the next logical step—change the design (perhaps requiring additional money and time) to get more TAR, or accept the system as-is, with its predicted level of military worth, and not invest more time and money.

10.3 The Step-by-Step Process

Participants in EW test and evaluation should follow a rigorous process that guarantees reliable results, insightful analysis, and useful information for the warfighter and decision makers. In particular, the EW test and evaluation process should follow the steps described in the following flowchart.

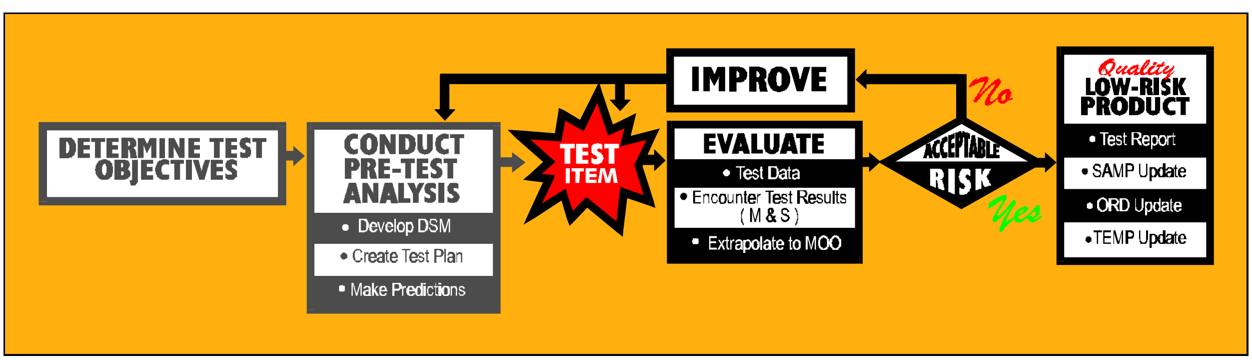

Figure 10-6. EW Test Process. *This flowchart describes the basic process we use to test and evaluate EW systems.*

The EW test and evaluation process includes the following steps:

- Determine test objectives
- Conduct pre-test analysis
- Test
- Evaluate
- Determine whether the system presents an acceptable risk
- Provide feedback to improve the system or test process
- Deliver a quality, low-risk system

10.3.1 Determine Test Objectives

For more information about early test and evaluation activities, see Section 5.3.1, Obtain Threat Scenarios for Modeling.

The goal of this step is to create clear and meaningful test objectives that specify what must be accomplished by test and evaluation. The objectives defined during this step should guide the test and evaluation process and facilitate thorough planning. In particular, the test objectives should enable us to create model and data sets that accurately reflect the threat, the platform or platforms, the environment, and the proposed solution.

As we have discussed throughout this report, the Partnership requires full involvement of all functional groups from the earliest stages of an acquisition. For the test and evaluation community, this early involvement is especially crucial. The test and evaluation community should work with the Integrated Concept Team (ICT) and Integrated Product Team (IPT) to develop test objectives that:

- Reflect warfighter needs.
- Identify the technical and operational issues that must be addressed.
- Identify possible areas of high risk.
- Specify the key performance indicators required by decision makers.
- Reflect mission, task, and performance requirements.

Test and evaluation personnel should participate in the entire process, beginning with the quantification of mission deficiencies and the establishment of requirements. Participation will allow these personnel to not only observe and gather information but provide guidance for the way requirements are specified. In this way, they can ensure that requirements are stated in terms that are meaningful to the warfighter and verifiable by our test and evaluation methods.

Account for All Important Test Factors

Assessing military worth demands a rigorous analysis of a number of variables, including:

- Modeling the threat system, including different factors such as operating modes, variants, and manufacturing variances
- Developing an understanding of how the EW system works
- Properly capturing all relevant data, including time and space positioning information
- Adequately characterizing the capabilities of the threat operator
- Estimating the reliability of the components
- Discerning how the reliability of theater intelligence can affect solution performance

These factors suggest a range of concerns and are not all inclusive. All relevant factors that can be modeled must be accounted for in the models we use to characterize the effectiveness and suitability of an EW solution.

Create Early Partnership Between Contractor and Government Test and Evaluation Communities

To ensure the success of these early planning efforts, government and industry must work together to create tests that provide the greatest insight into a developing system. In particular, government and industry share the responsibility for:

For more information about early test activities, see Chapters 5 and 6.

- Assessing the test infrastructure and making appropriate investment decisions
- Participating in foreign materiel exploitation (FME)
- Planning for test activities, including how to divide responsibilities for tests, evaluations, and training

The key to a successful test program, for both the government and industry, is early and open communication.

The earliest data-gathering activities are performed by the contractor. These activities allow the contractor to gain confidence as system development progresses. The government test community should be intimately involved in these contractor tests. Through this involvement, government can gain insight into the developing system and plan for later tests and evaluations.

Additionally, this cooperation between the government test community and the contractor will permit the government test community to gain early insight into the function and capabilities of the developing solution. When the developing contractor has a system ready for delivery, the government performs its own tests. The results of both contractor and government testing are used to support program decisions.

Another benefit of a closer working relationship between contractor test activities and government test activities is that we will minimize the duplication of effort and resources. Whenever possible, the government should encourage the contractor to use government test facilities and resources, so that the contractor does not build and equip facilities that unnecessarily duplicate government facilities. By the same token, the government should not recreate industry facilities. In every case, we should try to determine which path is most cost effective.

10.3.2 Conduct Pre-Test Analysis

During pre-test analysis, we make predictions about the test scenario, the test environment, and the effects of the proposed solution. More generally, test and evaluation personnel refine the test objectives established in the previous steps and determine what must be further defined and developed in order to accomplish those objectives.

During this step, we must determine whether current test and evaluation resources are sufficient for performing the proposed tests. For this reason, the pre-test analysis phase should begin as early as possible, so we arrange the necessary funding, acquire the required resources, and prepare the necessary facilities and models.

Conduct Foreign Materiel Exploitation with Silver Bullet Van

The effectiveness of an EW system depends on our accurate and thorough knowledge of threats. This knowledge is gained through a process known as foreign materiel exploitation (FME). Two new activities advocated by the Partnership should help us to make this activity more effective:

- Direct industry involvement in FME
- Use of the Silver Bullet Van

As a first step toward increasing industry involvement in FME, we will develop ways to widely and rapidly disseminate the results of DoD analyses of foreign weapon systems. The details for this process—which need to ensure that all interested and qualified contractors are given a chance to participate without compromising the security of data—are currently being determined by the government intelligence community.

After we have refined the methods for widely and rapidly disseminating FME data, we want to develop methods for allowing contractor representatives to participate directly in the process, by asking questions and making suggestions for improving the analysis. In the spirit of partnership, we feel that industry involvement is crucial—it can benefit all of us by allowing industry to gain greater insight into the kinds of threats their solutions are intended to counter.

Another new tool we will use to gather information about potential threats is the Silver Bullet program. This program conducts vulnerability assessments of threat weapon systems. Vulnerability assessments performed through the Silver Bullet program allow us to more accurately assess the impact of specific EW techniques and create models of threat systems that represent a higher level of fidelity.

10.3.3 Test

The goal of the test step is to gather data that can be used to evaluate a system's performance and its contribution to mission success. When we use the word test in this chapter, we mean any technical evaluation that involves a test item consisting of either hardware or software that will be used in the developing system.

Depending on the maturity of the system and the data needed for a particular evaluation, we employ different test methods. As we discussed in Section 10.2.2, Develop Greater Modeling and Simulation Capabilities, we need to use DSMs to achieve correlation between different types of testing and ensure that we get the appropriate information from each type to support technical and programmatic decisions.

Determine Military Worth

If military worth is quantified in terms of warfighter tasking—for example, placing targets at risk—we can use test data to make informed decisions about cost, schedule, and performance and increase the chances of warfighter mission success.

During the test step, we need to continually apply the Military Worth Method that guided us through earlier stages. Much of the data collected during developmental testing can help us assess the military worth of a developing solution. Data collected during operational testing, however, *must* clearly link to military worth and be checked against the warfighter's need.

Other data might also be collected to support test objectives, such as suitability or pilot ease-of-use, but the military worth pyramid should drive the test process. It is the means to link back to the requirement and forms the analytical framework that supports acquisition decisions.

A crucial difference in a military worth test and evaluation process is that results are used to improve the digital models we use throughout the acquisition and determine that the system is effective and suitable in an operational environment. The DSM allows us to make military worth assessments that are meaningful to the warfighter.

Assess RiO as a Function of Altitude and Aspect Angle

For more detail on the evaluation of military worth, see Chapter 4, Quantifying Military Worth.

All evaluations of the military worth of an EW system need to assess whether and how the system produces RiO against specific threats. To produce this assessment, we need to follow a process that captures all data necessary to characterize both wet and dry P_k grids with offset-based encounters.

In particular, tests need to follow these steps:

1. Assume the platform flies straight and level.
2. Generate engagements at particular cells in the P_k grid.
3. Calculate grid P_k (the probability the aircraft dies given it enters a grid cell relative to a specific threat).
4. Aggregate grid P_k s and calculate encounter P_s (the probability an aircraft lives given it enters a threat's lethal envelope).
5. Calculate the offset of the threat, where encounter P_s equals the attrition threshold.
6. Compare dry offset with wet offset and calculate RiO based on the difference.

By following this method, we can ascertain the effect of a system on particular threats. This information can then be used to understand whether and how the system contributes to mission success.

Correlate Tests with Actual Performance

Differences will always occur between the results of initial modeling and simulation of a solution, the testing of a prototype, and the real-world performance of the final product. Discrepancies will occur between how we believe the solution might behave, how we simulate it, and how it really functions in combat. These differences may occur for a variety of reasons and a great deal of effort goes into understanding the causes for these discrepancies.

Nevertheless, we can obtain results that show sufficiently close correlation so we can make acquisition decisions with enough confidence that the resulting system will meet the warfighter's needs. The key to understanding and accounting for the effects of these differences is continual correlation of test results with the models we use.

To take maximum advantage of modeling and simulation tools, we need to institute a rigorous feedback loop that ensures a clear correlation between the DSMs of the threat and the system under test and the results of test and evaluation activities. This feedback loop helps us to improve the models so we can make the best assessments of the military worth of the system under test throughout its development.

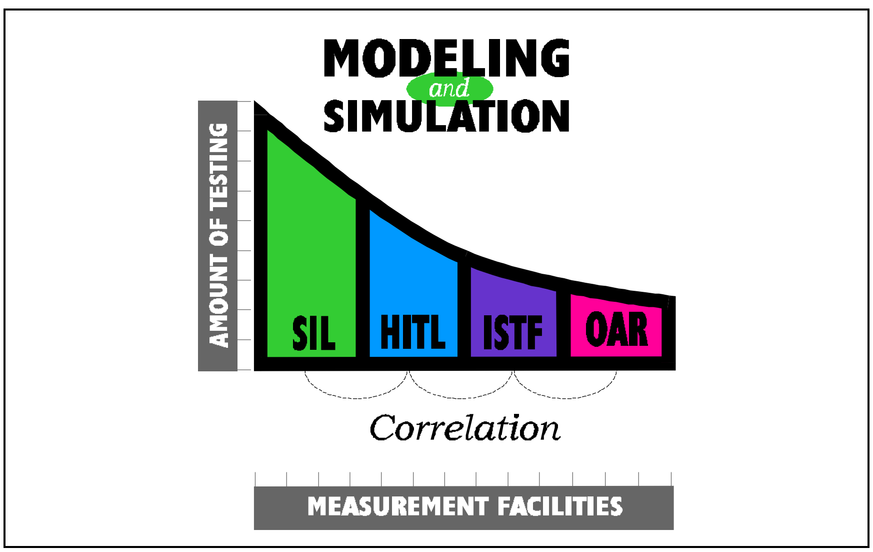

Figure 10-7. Correlation Between Models and Types of Tests. *We must ensure that results are correlated and a consistent feedback loop is established throughout the test and evaluation process.*

Correlating differences between model predictions and testing. Some factors that contribute to differences between model predictions and testing are:

- Threat characterizations
- Human factors as modeled on the range and in combat
- Test aircraft characterizations
- EW systems characterizations
- Actual versus predicted environmental conditions
- Background signal environments

Another way differences can occur is through analyzing the result with different methods, assumptions, and data. The key to mitigating these differences is to maintain a consistent modeling and simulation application based on an accepted DSM of the solution we are testing.

Correlating differences between the test environment and the scenario. Data from the test environment, whether early tests of system components or OAR testing, need to be correlated with the real-world scenario defined by the warfighter. Some of this correlation is achieved when the warfighter plans missions on the digitized version of the test range (which simulate the missions planned in the real-world scenario). This activity is described in Section 5.3.1, Obtain Threat Scenarios for Modeling.

Correlating data from different test facilities. Historically, the correlation of data from different test facilities has been poor. This lack of correlation can be a large problem if:

- Calibrations are off.
- Site operators have different expertise levels.
- Sites operate in different modes.
- The system under test is not designed to respond to threat variances.

Our use of digital modeling and simulation and detailed analysis will improve this type of correlation by identifying and eliminating fixed biases and point to facility calibrations and/or test modifications that will improve correlation. Test results from different facilities should then become complementary—both the amount of system knowledge and our associated confidence level will increase with each stage of testing.

10.3.4 Evaluate

While the test step gathers necessary data, we do not have the information necessary for making a decision until we make an evaluation of that data. Evaluation entails a thorough review of test data and relating that data to the predicted performance and operational capabilities of the system.

With the Partnership's Military Worth Method, evaluation requires assessing the system's contribution to the accomplishment of the mission objective, for example, putting targets at risk. In this way, any evaluation of a system is based on a quantifiable measure of whether the system solves a warfighter need.

The evaluation step should result in conclusions and recommendations that help decision makers understand how well or poorly the system satisfied the test objectives and how these relate to overall performance objectives.

In particular, these recommendations and conclusions should focus on:

- The ability of the system to accomplish its most critical task
- Implications of the test and evaluation results for the next planned step for the program

The evaluation should not be a pass/fail assessment of the system, but instead should provide insight for decision making.

Evaluation is not complete until all test data and objectives have been analyzed, and any differences between predicted and measured values have been resolved. For more information on how we respond to these deviations, see Section 10.3.6, Provide Feedback to Improve the System or Test Process.

10.3.5 Determine Whether the System Presents an Acceptable Risk

Two types of decisions need to be made after an evaluation is complete.

- First, we need to determine whether the test data is sufficient and reliable.
- Second, given adequate data, we need to decide whether the system is mature enough (that is, suitable and effective enough given the phase of the acquisition) to proceed in the acquisition process.

In the final analysis, the program manager or a higher acquisition official must decide whether the system presents an acceptable risk and whether changes are necessary. To this end, the program manager needs the test manager to:

- Determine whether the test demonstrated the objectives
- Ensure the test design was adequate for measuring the effectiveness of the system
- Make a recommendation on the maturity of the system

10.3.6 Provide Feedback to Improve the System or Test Process

It is imperative that we learn from the test and evaluation process and strive to continually improve the performance of our solutions, the accuracy of our tests, and the fidelity of our models. This continual improvement demands a feedback process that captures what we learn and provides guidance for the next round of tests or development.

Any time we leave a test facility, we should be able to enhance and modify our mathematical understanding of how well we think our solution performs. In particular, when a system matriculates into the next phase, the data collected during the testing of the previous phase must be used to support subsequent phases of an acquisition.

Before proceeding with the improvement step, we must determine which aspect or aspects require modification. Any discrepancy between our expected result and the test result should cause us to look for the appropriate explanation. We should not, for example, move ahead with changes to the solution until we have ruled out possible inaccuracies in the model, since it’s always less expensive to fix the model than the actual solution.

Providing feedback for improvement is discussed in the following order:

- Improve the DSM
- Improve the test method
- Improve the system

Improve the DSM

As discussed in Section 10.2.2, DSMs are the basis for correlating the results of various types of tests, and so the Partnership puts high value on developing and improving DSMs. Consequently, we place a high priority on assessing their accuracy and using feedback from tests to improve their fidelity.

Our greater reliance on digital models will allow us to increase the cost-effectiveness of testing and create a reliable process for tracing the military worth of a solution from technical attributes up through mission and campaign objectives.

If any discrepancy between expected and actual results is caused by a mathematical inaccuracy in the model, we need to change the model.

The process for using models and simulations during test and evaluation should follow the following steps:

1. Create or update the model.
2. Make a prediction of the result.
3. Perform the test.
4. Compare the prediction to the result.

This process should be repeated as often as we collect new test data. The accuracy of models depends on complete and precise data about a large number of variables, and we can constantly improve them by updating them with the results from laboratory or range tests. Particularly, actual test results allow us to refine assumptions we made earlier in the acquisition process.

The premise of DSMs is that if we really understand a threat and our attempt to create solutions to that threat, we can describe them mathematically and build models of them. These models depend on assumptions that should be verified and validated in actual tests. Any discrepancy should be seen as an opportunity to refine the assumptions that were originally made during the model's creation or refinement.

While the accuracy of lower-level models and test processes is enhanced by data gathered in higher-fidelity tests, the reverse is also true.

Improve the Test Method

Occasionally, the difference between expected and actual results is caused by the test method. During the evaluation of the system, we must also evaluate the test methods we use. Any problems that arise could have significant effects on later stages of the acquisition cycle. For example, failure to respond to feedback information during developmental testing (DT) can adversely affect the outcome of operational testing (OT) and put a program's success in jeopardy. Therefore, any problems with test methods should be addressed as quickly as possible.

Improve the System

If a system falls short on some particular performance specification on its own merits, we do not automatically return to development to boost its performance. Instead, we follow these three steps:

1. Determine the cause of the problem.
2. Assess its effects on military worth.
3. Decide whether the effect on military worth, if any, justifies additional investment of time and money.

If the deviation detected by test and evaluation affects the military worth of the system, its schedule, or its cost, we need to go back to the warfighter to determine whether the solution is still worth pursuing. If, however, the deviation does not affect any of these three dimensions, the program manager can decide whether any response is necessary.

This approach ensures that we do not blindly—and expensively—try to meet a specification for its own sake, but instead assess the effect of a discrepancy on our attempt to satisfy the warfighter's requirement.

10.3.7 Deliver a Quality, Low-Risk System

When the EW test process is followed and Military Worth Method is consistently applied, we can reduce risk and ensure that a solution is mature enough to advance in the acquisition cycle. Test and evaluation processes that are well-designed and well-executed provide the warfighter with confidence in the results and with insight into the operation of a solution.

Our ultimate goal is always providing the warfighter with solutions to deficiencies. Following a disciplined approach to test and evaluation helps us to guarantee that the systems we develop are delivered to the warfighter in a better, faster, and cheaper manner. The test and evaluation process we have described here provides maximum insight for decision makers, allowing them to make the best decisions within the requirements trade space and provide the best value for the warfighter.

Summary

This chapter discussed the Partnership approach to evaluating the results of acquisition efforts. By instituting a disciplined process for evaluation, we can ensure that we link the results of our efforts to the warfighter's deficiencies. In addition, we can make insightful decisions about the value of the systems we develop.

The analysis and insight gained from evaluating the result in this way provides the foundation for the next stage of the acquisition. At the end of each phase, the acquisition faces a milestone decision. Presuming the system passes this decision, we return to the set of activities described in Chapter 7, Convey the Requirements to begin the next phase.